

KATHERINE L. ELKINS, Ph.D.

Professor of Humanities & Comparative Literature · Kenyon College
Director, Integrated Program in Humane Studies: AI, Humanities, and Society
Founding Co-Director, AI CoLab · Co-Founder, Human-Centered AI Lab
elkinsk@kenyon.edu · katherineelkins.com
ORCID 0000-0001-9887-4854 · en.wikipedia.org/wiki/Katherine_Elkins

PROFILE

Professor and researcher working at the intersection of literature, philosophy, and artificial intelligence. Author of *The Shapes of Stories* (Cambridge University Press) and editor of *Proust's In Search of Lost Time: Philosophical Perspectives* (Oxford University Press), with wide-ranging scholarship on narrative, memory, perception, and the philosophy of mind. Co-creator of the first human-centered AI curriculum (2016) and an early empirical investigator of large language models, combining humanistic interpretation with hands-on computational and AI-safety research. Field-defining contributions include the first ensemble-based method for diachronic sentiment analysis of narrative, the first study testing whether an AI system could write convincingly as a human author (GPT-3, 2020), and the first comparative ethics-based audit of deployed LLMs. Leads the Modern Language Association's team at the US AI Safety Institute (now NIST CAISI) and serves as Principal Investigator on Schmidt Sciences' Archival Intelligence initiative.

AREAS OF SPECIALIZATION

Specialization: AI and the humanities; narrative theory and computational narratology; philosophy and literature; philosophy of mind, cognition, and neuroscience; aesthetics; information theory and meaning; computational and digital humanities; AI ethics, safety, and governance.

Competence: translation studies; memory studies; modernism; comparative literature; lyric and poetics; creativity and authorship studies.

EDUCATION

Ph.D., Comparative Literature. University of California, Berkeley, 2002.

B.A., Literature (Distinction in Major, cum laude). Yale University, 1990.

ACADEMIC APPOINTMENTS

Professor of Humanities and Comparative Literature, Kenyon College (2018–Present).

NEH Distinguished Teaching Professor of Humanities and Comparative Literature, Kenyon College (2020–2023).

Associate Professor, Kenyon College (2010–2018).

Assistant Professor, Kenyon College (2002–2010).

ADMINISTRATIVE & PROGRAM LEADERSHIP

Director, Integrated Program in Humane Studies: AI, Humanities, and Society, Kenyon College (2020–Present).

Founding Co-Director, AI CoLab (formerly KDH AI Collaboratory), with Jon Chun, Kenyon College (2017–Present).

Founding Director, Comparative World Literature Program, Kenyon College (2012–2020, with sabbatical leave).

Co-Founder and Co-Director, Human-Centered AI Lab, Inc. (Ohio nonprofit), with Jon Chun (2024–Present).

BOOKS & EDITED VOLUMES

The Shapes of Stories: Sentiment Analysis for Narrative. Cambridge Elements in Digital Literary Studies. Cambridge University Press, 2022.

Proust's In Search of Lost Time: Philosophical Perspectives. Oxford Studies in Philosophy and Literature. Oxford University Press, 2022. (Editor and Contributor; includes the volume Introduction and the chapter "Proust's Consciousness.")

Guest Editor, "Depiction of Good and Evil in Fairytales," *Humanities* (MDPI), special issue, 2024–25.

PUBLICATIONS

Under Review & In Progress

"When Prohibitions Become Permissions: Auditing Negation Sensitivity in Language Models" (with J. Chun). arXiv:2601.21433, 2026. Under review.

"Biased or Just Noisy? Why Aggregate AI Audits Miss Scenario-Specific Vulnerability" (with J. Chun). Under review, 2026.

"The AI Governance Trilemma" (with C. Schroeder de Witt and J. Chun). Under review, 2026.

"The AI Fiction Paradox." Preprint, arXiv:2603.13545, 2026.

Forthcoming & Accepted

"Position: If Open Source Is to Win, It Must Go Public" (with J. Z. Tan, N. Vincent, M. Sahlgren, J. Low, D. Pham, S. Pyysalo, and J. Jitsev). *Proceedings of ICML 2026* (accepted, Position Paper track). Earlier spotlight version, CODEML Workshop, ICML 2025 (arXiv:2507.09296).

"MADCal: Multi-Agent Debate Calibration" (with J. Chun). NeurIPS 2026.

"The Paradox of Robustness: Decoupling Rule-Based Logic from Affective Noise in High-Stakes Decision-Making" (with J. Chun). arXiv:2601.21439, 2026. NeurIPS 2026.

"AgenticSimLaw: A Juvenile Courtroom Multi-Agent Debate Simulation for Explainable High-Stakes Tabular Decision Making" (with J. Chun and Y. S. Lee). AAI-26 LaMAS, Singapore, January 2026.

"Can We Fall in Love with AI Fiction? The AI Fiction Paradox." *Modern Fiction Studies*, Special Issue on Cultural AI. Forthcoming, Summer 2027.

"Storytelling with Language Models: Narrative Theory Meets Generative AI." *Narrative Inquiry*. Forthcoming.

"The Dual Edge: AI's Impact on Cultural Heritage and Preservation." UNESCO. Forthcoming.

AI, Creativity & Authorship

"Can GPT-3 Pass a Writer's Turing Test?" (with J. Chun). *Journal of Cultural Analytics* 5, no. 2, 2020.

"AI Comes for the Author." *Poetics Today* 45, no. 2, 2024, pp. 267–274.

"A(I) University in Ruins: What Remains in a World with Large Language Models?" *PMLA* 139, no. 3 (special Theories and Methodologies issue), 2024, pp. 559–565.

Narrative & Computational Methods

"What the Rise of AI Means for Narrative Studies" (with J. Chun). *Narrative* 30, no. 1, 2022, pp. 104–113.

"Beyond Plot: How Sentiment Analysis Reshapes Our Understanding of Narrative Structure." *Journal of Cultural Analytics* 10, no. 3, September 2025.

- “Storytelling with Language Models: Narrative Theory Meets Generative AI.” *Narrative Inquiry*. Forthcoming. [See Forthcoming & Accepted.]
- “The Shapes of Cinderella: Emotional Architecture and the Language of Moral Difference.” *Humanities*, special issue on fairy tales (Guest Editor), 2025.
- “eXplainable AI with GPT-4 for Story Analysis and Generation: A Novel Framework for Diachronic Sentiment Analysis” (with J. Chun). *International Journal of Digital Humanities* 5, no. 2–3, 2023, pp. 507–532.
- “In Search of a Translator: Using AI to Evaluate What’s Lost in Translation.” *Frontiers in Computer Science* 6:1444021, 2024.
- “Can Sentiment Analysis Reveal Structure in a ‘Plotless’ Novel?” arXiv, 2019.

AI Safety, Ethics & Governance

- “Position: Near to Mid-term Risks and Opportunities of Open-Source Generative AI.” *Proceedings of ICML 2024*. Oral presentation (top ~2%). Operationalizes an AI-openness taxonomy across forty large language models, with an international author team (Oxford Foerster Lab and Torr Vision Group, UC Berkeley, Notre Dame, Berklee, Luxembourg Institute of Science and Technology, University of Luxembourg, University of Bath, Tec de Monterrey, National University Philippines, and others). arXiv:2404.17047.
- “Position: If Open Source Is to Win, It Must Go Public” (with J. Z. Tan, N. Vincent, M. Sahlgren, J. Low, D. Pham, S. Pyysalo, and J. Jitsev). *Proceedings of ICML 2026* (Position Paper track); earlier spotlight version, CODEML Workshop, ICML 2025 (arXiv:2507.09296). [See Forthcoming & Accepted.]
- “Comparative Global AI Regulation: Policy Perspectives from the EU, China, and the US” (with C. Schroeder de Witt and J. Chun). arXiv:2410.21279; SSRN 5104429, 2024.
- “Provenance Infrastructure as a Safeguard for Cultural Commons in the Age of Generative AI” (with J. Chun, Human-Centered AI Lab). Selected contribution (1 of 23 from 43), UNESCO Subgroup on Intellectual Property and Culture, 2026.
- “Informed AI Regulation: Comparing the Ethical Frameworks of Leading LLM Chatbots Using an Ethics-Based Audit” (with J. Chun). arXiv:2402.01651, 2024.
- “Syntactic Framing Fragility: An Audit of Robustness in LLM Ethical Decisions” (with J. Chun). arXiv:2601.09724, 2025. Submitted, FAcCT 2026.

Philosophy, Mind & Literature

- “Naming the Lyric: Literature versus Philosophy in Plato’s Symposium.” *Philosophy and Literature* 44, no. 2, 2020, pp. 402–417.
- “On Kafka” (“Kafka’s Picture Lessons in The Trial”). *MLN* 136, no. 3 (German issue), 2021, pp. 755–769.
- “Memory, Technology and Wisdom.” *Graduate Faculty Philosophy Journal*, 2022.
- “Memory and Material Significance: Composing Modernist Influence.” *Modern Language Quarterly* 69, no. 4, 2008, pp. 509–531.
- “Middling Memories and Dreams of Oblivion: Configurations of a Non-Archival Memory in Baudelaire and Proust.” *Discourse* 24, no. 3, 2002.
- “Stalled Flight: Horatian Remains in Baudelaire’s ‘Le Cygne.’” *Comparative Literature Studies* 39, no. 1, 2002. (A. Owen Aldridge Prize.)
- “History’s Theft and Memory’s Return in Maryse Condé.” 2006.
- “Proust’s Family at Home.” *H-France Salon* 13 (Proust at Home), 2021, pp. 1–9.

Chapters in Edited Volumes

“Generative AI and the Stakes of Literary Form,” in *The New Literature and AI Studies*, ed. Robert Marzec. Cambridge University Press. Forthcoming. (Volume contributors include N. Katherine Hayles, Matthew Kirschenbaum, Alan Liu, Rita Raley, and Lauren Goodlad.)

“Proust’s Novel Time,” in *Time: A Multidisciplinary Introduction*, eds. Kattan Gribetz and Kaye. De Gruyter Oldenbourg, 2023, ch. 7.

“Wordsworth’s Literary Sublime,” in *Fictional Worlds and Philosophical Reflection*, ed. Garry Hagberg. Springer / Palgrave Macmillan, 2022, pp. 227–244.

“No Return: Memory as Exile in Maryse Condé,” in *Maryse Condé: By Way of Introduction*, ed. Barbour. Africa World Press, 2006.

AI, Pedagogy & Society

“The Crisis of Artificial Intelligence: A New Digital Humanities Curriculum for Human-Centred AI” (with J. Chun). *International Journal of Humanities and Arts Computing* 17, no. 2, 2023, pp. 147–167. The article that defined what a human-centered AI curriculum looks like; coined the term “AI Digital Humanities (AI DH).”

“The Evolution of AI” (with J. Chun). Edinburgh University Press blog, 2023.

Additional humanities scholarship appears in *PMLA*, *Poetics Today*, *MLN*, *Discourse*, *Modern Language Quarterly*, *Philosophy and Literature*, *Comparative Literature Studies*, *Narrative*, and the *Graduate Faculty Philosophy Journal*.

Reviews

Review of *Proust as Philosopher: The Art of Metaphor* by Miguel de Beistegui. *Notre Dame Philosophical Reviews*, 2015.

Review of *La Littérature d’imagination scientifique* by Daniel Fondanèche. *The French Review* 88, no. 3, 2015, pp. 211–212.

Review of *Introduction à la méthode posttextuelle: l’exemple proustien* by Franc Schuerewegen. *The French Review* 87, no. 3, 2014, p. 279.

Review of *Never Say I: Sexuality and the First Person in Colette, Gide, and Proust* by Michael Lucey. *Modernism/modernity* 17, no. 4, 2010.

Review of *The Re-Enchantment of the World: Secular Magic in a Rational Age*, ed. Joshua Landy and Michael Saler. *Modernism/modernity* 15, no. 3, 2010, pp. 963–964.

Audio Lecture Series (Public Scholarship)

The Modern Novel. The Modern Scholar Series, Recorded Books / Audible, 2013.

The Giants of French Literature: Balzac, Flaubert, Proust, and Camus. The Modern Scholar Series, Recorded Books / Audible, 2010. Reviewed in *Library Journal* as “a compelling survey course.”

Odyssey of the West, Part VI: Towards the Modern. The Modern Scholar Series, Recorded Books / Audible, 2009.

SELECTED SCHOLARLY RECEPTION

Mentored student research has been downloaded more than 107,000 times, drawn from 400+ projects and reaching 198 countries. Across published work, Google Scholar records roughly 794 citations (February 2026); the most-cited works include “Can GPT-3 Pass a Writer’s Turing Test?,” “Risks and Opportunities of Open-Source Generative AI,” and *The Shapes of Stories*. The named reception below records specific citers and structural adoption across the humanities, NLP, and AI venues.

Field-defining contributions include: first human-centered AI curriculum (2016); first ensemble-based method for diachronic sentiment analysis of narrative; first study testing whether an AI system could

write convincingly as a human author (GPT-3, 2020); first comparative ethics-based audit of deployed LLMs; first operational Syntactic Framing Fragility / Syntactic Variation Index method.

Computational Humanities & Narrative

The Shapes of Stories — the ensemble methodology for diachronic sentiment analysis adopted as evaluation infrastructure (Hamilton, Wilkens, and Piper, NarraBench, EACL 2026), extended in visual analytics (Yeh et al., Story Ribbons, IEEE TVCG 2026), applied across 25,728 narrative retellings (*Scientific Reports*, 2023), and used to analyze literature, translations, TV scripts, end-of-life medical narratives, and social-media narratives of elections and economic crises. Taken up across the Cambridge Elements in Digital Literary Studies series.

“Can GPT-3 Pass a Writer’s Turing Test?” — published months after the GPT-3 release and archived on gwern.net, a technical repository consulted by researchers at organizations including OpenAI and DeepMind; cited by Floridi and Chiriatti (*Minds and Machines*, 2020); the “writer’s Turing test” adopted as a named framework across philosophy of AI and NLP evaluation, and engaged in the machine-psychology (Hagendorff et al.) and AI Ghostwriter (ACM TOCHI) literatures.

“The Crisis of Artificial Intelligence: A New Digital Humanities Curriculum for Human-Centred AI” (IJHAC, 2023) — the article that defined what a human-centered AI curriculum looks like and coined “AI Digital Humanities (AI DH)”; cited by Terence Tao (Fields Medal) and Tanya Klowden in “Mathematical Methods and Human Thought in the Age of AI” (forthcoming, *Blackwell Companion to the Philosophy of Mathematics*; arXiv:2603.26524), and adopted across AI ethics, education, information science, and digital humanities.

“eXplainable AI with GPT-4 for Diachronic Sentiment Analysis” (IJHAC, 2023) — cited by Cugurullo and Xu (*Policy and Society*, 2024).

“Beyond Plot: How Sentiment Analysis Reshapes Our Understanding of Narrative Structure” (*Journal of Cultural Analytics*, 2025) — selected as a Digital Humanities Now Editors’ Choice (2025).

AI, Authorship & the University

“AI Comes for the Author” (*Poetics Today* 45.2, 2024) — engaged by James Phelan (*Poetics Today*, 2024) as part of an instructive debate.

“A(I) University in Ruins” (*PMLA* 139.3, 2024) — anchors a central theoretical claim in Cao et al. (2025).

AI Ethics & Governance

“Informed AI Regulation” — the confidence-score audit applied across 1,613 decision scenarios (Liu, Liu, and Yu, COLING 2025) and named among foundational works in an AAAI 2026 alignment survey.

Open-source position paper (ICML 2024 oral) — named among canonical openness frameworks (Paris, Moon, and Guo, FAccT 2025).

Literary Scholarship: Memory, Lyric & Interpretation

“Stalled Flight: Horatian Remains in Baudelaire’s *Le Cygne*” (*CLS* 39.1, 2002; A. Owen Aldridge Prize) — cited as canonical authority on Horace 2.20 as the model for *Le Cygne* (Singh, *Comparative Literature*, 2012; Marinčič, *Acta Neophilologica*, 2009).

“Middling Memories and Dreams of Oblivion” (*Discourse* 24.3, 2002) — the non-archival memory framework adopted in the Van Hulle / Antwerp genetic-criticism ecosystem (Luyten, 2015; Johnston, *Critique*, 2018).

“Naming the Lyric” (*Philosophy and Literature* 44.2, 2020) — cited in Rick Anthony Furtak, *Kierkegaard, Socrates, and the Meaning of Life* (Cambridge University Press, 2025).

Proust’s In Search of Lost Time: Philosophical Perspectives (Oxford University Press, 2022) — engaged in Tom Stern (*European Journal of Philosophy*, 2025).

GRANTS, FELLOWSHIPS & AWARDS

Research Grants & Funded Projects

Schmidt Sciences HAVI Award (2025). Principal Investigator, “Archival Intelligence: Rescuing New Orleans’ Endangered Cultural Legacy.” One of 23 awards worldwide (from 600 applications); up to \$330,000 over 18 months within an \$11M program. Develops open-access AI tools (including smartphone-based capture) to preserve endangered community archives, with a pilot on multilingual Creole and Cajun newspapers and early jazz artifacts, addressing “cultural flattening” in AI models.

Notre Dame–IBM Tech Ethics Lab Grant (2024). PI (\$60,000), “How Well Can GenAI Predict Human Behavior? Auditing LLMs for Fairness, Accuracy, Transparency, and Explainability (FATE),” with Jon Chun and Yong Suk Lee. Uses juvenile-recidivism prediction as a test case for multi-agent decision support; produced the AgenticSimLaw framework.

Fellowships

Whiting Fellowship (Mrs. Giles Whiting Foundation) — research in comparative literature and the philosophy of literature.

Whiting Summer Research Award.

Andrew Mellon Fellowship — graduate research.

Genevieve McEnerney Fellowship, University of California, Berkeley.

German Academic Exchange Service (DAAD) Fellowship, University of Regensburg.

University Fellowships, University of California, Berkeley; Aix-en-Provence.

John S. Koch Memorial Fellowship — Oxford summer study.

Honors, Prizes & Teaching Awards

NEH Distinguished Teaching Professorship, Kenyon College (2020–2023).

A. Owen Aldridge Prize in Comparative Literature (2001), for an essay on Charles Baudelaire.

Senior Faculty Trustee Teaching Excellence Award, Kenyon College (2023).

Baccalaureate Address Speaker, selected by the senior class, Kenyon College (2024).

Teaching Effectiveness Award, University of California, Berkeley.

Outstanding Graduate Student Instructor Award, University of California, Berkeley.

KEYNOTES & DISTINGUISHED LECTURES

“Governing AI in Partial View: Why Responsible AI Needs the Humanities.” Keynote, University of Michigan MIDAS Annual Data Science and AI Summit (December 2026, accepted).

“Why Can’t We Fall in Love with AI Fiction?” International Society for the Study of Narrative, Aarhus (June 2026).

“Leveraging AI for Research.” OpenAI Inaugural Higher Education Forum, San Francisco (October 2025). Selected as one of only 6 faculty speakers worldwide from 1,000 applicants.

“The Power of Connection: Leveraging Technology for Humanistic Medical Education.” Weill Cornell Medicine–Qatar, Doha, METC Conference (October 2025). Continuing medical education credits.

“Responsibly Engaging in Generative AI.” Faith, Reason and World Affairs Symposium, Opening Plenary, Concordia College, Moorhead, MN (September 2025).

“Artificial Intelligence and Cultural/Creative Industry.” UNESCO MONDIACULT Conference, Cairo (April 2025).

“Tomorrow’s Agents, Today’s Questions: Emotional Intelligence and Human Identity at the Dawn of AI Networks.” Kahn Liberal Arts Institute, Smith College (February 2025).

“Human Algorithms.” RALLY Innovation Conference, Indianapolis (2025).

Baccalaureate Address (nominated by graduating seniors): “Revisiting David Foster Wallace’s ‘This Is Water’ 20 Years Out.” Kenyon College (December 2024).

“Beyond AI Literacy.” Dartmouth (2024).

“Can AI Reason Ethically?” Northrup Distinguished Lecture, Yale University (February 2024).

“The Value of the Liberal Arts in the Age of AI.” The A. J. Carlson Lecture, Austin College (2024).

“AI Literacy Across the Curriculum.” Lafayette College (2024).

“Human-Centered AI: High-Impact Change from the Classroom to the Lab.” Carleton College Day of Digital Humanities (May 2024).

“What is a Story? Lessons from Nine Years of AI Modeling.” Washington University in St. Louis, “Stories that Win” Symposium, Incubator for Transdisciplinary Futures (2024).

Inaugural Plenary, NIST AI Safety Institute Consortium (AISIC) (2024).

“Affective AI.” Keynote Address, Wofford College (October 2023).

“The AI Epoch: Ethics and Insights.” The Meredith-Donovan Lecture, Mount St. Mary’s University (2023).

“Like the Invention of Fire: The Dangers and Opportunities of Generative AI.” Denison University (2023).

“Generative AI: Automating Intelligence.” Keynote, Virginia Wesleyan University (2023).

“Programming Humanity” / “Comparative Studies: Imagining Futures in the Age of Information.” Keynote, Ohio State University (April 2019).

CONFERENCE PRESENTATIONS & INVITED TALKS

2026–2027

“AI and Fiction,” panel (Prose Fiction forum), with Tess McNulty and Teddy Roland. MLA Convention, Los Angeles (January 2027).

Organizer and respondent, Human-Centered AI Lab panel “What Constructed Agents Sound Like: Language, Pragmatics, and Culture in Human–AI Interaction.” TSL 2026 (Technology for Second Language Learning), Iowa State (October 2026).

Invited participant, Schmidt Sciences HAVI 2026 Convening, London (October 2026).

“Civic Education After AI” / AI and Democracy. Ohio State University, Chase Center (April 2026).

Co-organizer, Roundtable on AI and Medicine. The Helix Center, New York (April 2026).

“Why Can’t AI Write a Bestseller?” International Society for the Study of Narrative (June 2026).

“Neuroscience and Literature.” Modern Language Association (January 2026).

2025

Co-host, “AI and Academic Research,” with Kevin Weil (VP, OpenAI for Science) and Leonardo Impett. OpenAI Forum (December 2025).

“EQ, Emotion, and AI: Including Work on Neurodivergence.” Weill Cornell Medicine–Qatar Med Tech Conference, Doha (October 2025).

“Can We Fall in Love with AI Fiction: The AI Fiction Paradox” / “Cultural AI: Thinking Artificial Intelligence Across the Disciplines.” Conference on Artificial Intelligence, Modern Fiction Studies Conference, Purdue University (September 2025).

“Computational Approaches to Narrative.” International Society for the Study of Narrative (April 2025).
“Data Humanities.” University of Colorado Boulder (March 2025).
“AI and Health: Navigating Emotion, Ethics, and Agency.” Teaching Prevention (March 2025).
“AI at a Crossroads: Who Shapes the Future?” Artist-in-Residence talk, WPI Global Lab, Worcester Polytechnic Institute (February 2025).
“AI for Social Good.” Yale University (February 2025).
Yale Alumni in AI (2025).
Presented Notre Dame–IBM grant findings (juvenile-recidivism multi-agent simulation), RISE AI & Society Conference, University of Notre Dame (January 2025).
“Rereading GenAI.” MLA, New Orleans (January 2025).
“Speaking the Language of Skills, Competencies, and Literacies in the Age of AI.” MLA MAPS Leadership Institute (department chairs workshop), MLA (June 2025), with John Baskerville (USMA), Todd Butler (San Diego State), and Abigail Heiniger (Lincoln Memorial).
“New Academic Programs for an AI-Driven Work Force.” Chronicle of Higher Education Virtual Forum (June 2025).

2024

“Emotional Hacking of LLMs.” US AI Safety Institute (December 2024). Presented experimental protocol demonstrating vulnerability to emotional manipulation.
“Ethical Auditing and Syntactic Framing.” OAIRC, Meta London (October 2024). Contributed to internal guidelines on model assessment.
“Rewriting Literary History with Algorithms.” University of Illinois Chicago (2024).
“Lost in Translation.” SlatorCon, Translation Intelligence Conference (2024).
“AI and What it Means to Create.” MLA (2024).
“Revisiting ‘The Trouble with Bias’ with GenAI.” MLA (2024).
“Literature, Language and Authorship in the Age of AI.” MLA (2024).
“How AI is Defining the Future of Organizations,” conversation with Sue Cantrell (Deloitte). HumanCentric Labs Webinar (2024).

2023

“Stories Have Shapes. Now What?” Computational Formalism, Dartmouth Neukom Institute (2023).
“The Shapes of Stories in the Public Sphere.” International Society for the Study of Narrative (2023).
“AI Needs the Humanities (and the Humanities Need AI),” with Matthew Kirschenbaum, Rita Raley, and Zoe LeBlanc. University of Tennessee, Knoxville (Humanities Center & Libraries) (2023).
“Human-Machine Decision Making and the Convergence of Intelligence” / “Computation May Be All You Need: Is Intelligence Convergent?” McGill Centre for Convergence / Pittsburgh Supercomputing Center / Carnegie Mellon, Precision Convergence Webinar (2023).
Featured lecture, AI Working Group, Wofford College (2023).
Invited lecture, Concordia University Montreal (2023).
“Every Institution Needs an AI Strategy.” Academic Leadership Conference (2023).
“Generative AI and the Future of Work.” HumanCentric Labs (2023).
“Generative AI: Challenges and Opportunities.” Reed College (2023).
“AI and How It’s Transforming the Workplace.” Academic Search (2023).
“AI and How It’s Transforming Higher Ed.” Academic Search (2023).

"AI Explained: Past, Present, and Future." Academic Search (2023).

2022

Helix Center Roundtable, "Coding and the New Human Phenotype: Are Natural Language Generators for Real?" with Ned Block (NYU), Francesca Rossi (IBM), and Kyunghyun Cho (October 2022).

Helix Center Roundtables, New York — with Ned Block, Francesca Rossi, Joseph LeDoux, Rosalind Picard, and Kyunghyun Cho (2022–23).

2018–2021

"Can AI Be Creative?" MLA (2021).

"How Narrative Feels." MLA (2021).

"Storytelling with AI." Ohio State University (2020).

"What Sentiment Analysis Can and Can't Tell Us Over Time." International Conference on Narrative (2020).

"In Search of Proust's Change Over Time." Proust and the Prix Goncourt, MLA, Seattle (January 2020).

"Reparative Narratives: Computational Approaches to Affect in Woolf." Digital Exhibit, Modernist Studies Association, Toronto (October 2019).

"Modernism's Sentimental Journeys." Modernist Studies Association, Toronto (October 2019).

First transdisciplinary AI research presentation, Modernist Studies Association (2019).

"AI and Comparative Studies." Ohio State University (2019).

"No Ordinary Plot." Modernist Studies Working Group (October 2018).

"Music and Narrative: Proust's *À la Recherche*." International Conference on Narrative, McGill University (April 2018).

Chair, "Memory, Memories, and Relative Time." *Metaphors of Time: An Interdisciplinary Conversation*, Ohio State University (April 2018).

"Proust's Japanese Balls: japonisme in Proust's *À la recherche*." *East-West Intertextuality in Modern Literature and Cinema*, ACLA, UCLA (March 2018).

2000–2017

"What is Digital Humanities 3.0?" Center for Innovative Pedagogy, Kenyon College (November 2017).

"Drawing Space: Rethinking Proust's Impression." ACLA, Utrecht (July 2017).

"Dark Matter," Chapter 2. Modernist Studies Group, Ohio State University (September 2015).

"Proust." Modernist Studies Group, Ohio State University (January 2011).

"Baudelaire's Acoustic Inspiration." Nineteenth-Century French Studies Conference, Yale University (October 2010).

"Vanishing Limits and the Modern: Proust's Dissolve." 20th and 21st Century French and Francophone Studies Colloquium, Washington, DC (March 2008).

"Proust's Exceptional Memory." 20th and 21st Century French and Francophone Studies Colloquium, College Station, TX (March 2007).

"Fragments of Memory: Proust's Characters as Non-Archival Markers." *Cognitive Approaches to Literature*, MLA, Philadelphia (2006).

"Literature as Cultural Philosophy: Proust as a Case Study." *Philosophical Approaches to Literature*, MLA, Philadelphia (2006).

"Sounding the Archive: Listening to Music in Proust." *Listening to Modernism*, Modernist Studies Association, Tulsa (2006).

“Of Script and Sound: Rethinking Typography in Proust.” Thinking Through Memory, UC Berkeley and Stanford Joint Conference (April 2002).

“(En)countering the Hexagon: Fiction versus Memoir in Maryse Condé.” Doreen Townsend Center for the Humanities (March 2002).

“History’s Theft, Memory’s Return in Condé’s Nanna-Ya.” MLA, Slavery, History and Memory (December 2001).

“No Return: Memory as Exile in Maryse Condé.” NYU French Department Colloque, Aliénation et exil (April 2001).

“Musical Limits: Proust’s Musical Memory as a Limit of Expression.” PAMLA, UCLA (November 2000).

POLICY, GOVERNANCE & ADVISORY ROLES

Lead, Modern Language Association team, US AI Safety Institute / NIST AI Consortium (now CAISI) (2024–Present). Consortium team: Jon Chun (co-lead), Rita Raley, Seth Perlow, Anna Mills.

Principal Investigator, Schmidt Sciences Humanities and AI Virtual Institute (HAVI), “Archival Intelligence: Rescuing New Orleans’ Endangered Cultural Legacy” (2025–Present).

Member, NIST Public Working Group on Generative AI (2023). Contributed to the AI 600-1 Generative AI Profile risk framework.

UNESCO AI & Cultural Heritage Initiative (2024–Present); UNESCO Subgroup on Intellectual Property and Culture, CSO Network on the Ethics of AI (2026).

Member, Public AI — international network for AI as public infrastructure (2023–Present).

OpenAI Higher Education Forum — Education Guild selected speaker (1 of 6; San Francisco, 2025).

Lead, Transparency Working Group, Meta Open Innovation AI Research Community (2023–2025).

Member, Global Policy Working Group.

Steering Committee, ADHO Computational Literary Studies (2026).

MLA AI & Research Working Group; MLA MAPS Leadership Institute (2025).

Advisory Board, The Helix Center: An Unhurried Search for Wisdom (2022–Present).

Former Executive Board Member and Chief AI Officer, HumanCentric Labs (2023).

Member, Women in AI; AI in Education.

STANDARDS & POLICY CONTRIBUTIONS

As lead of the Modern Language Association team in the NIST AI Consortium (CAISI) since its 2024 founding, authored three public comments to NIST in 2026, each on a different dimension of how AI systems are secured and tested: the Agent Security RFI (docket NIST-2025-0035), which introduced the concept of interpretive tractability; the AI 800-2 benchmarking framework; and the NCCoE concept paper on AI agent identity and authorization. Authored an MLA newsletter feature (2026) introducing the consortium’s NIST work and the concept of interpretive tractability to a humanities readership.

TEACHING & CURRICULUM INNOVATION

Co-created the first human-centered AI curriculum (2016) and co-founded the AI CoLab with Jon Chun, building a research-driven program whose mentored student work has reached a global audience. The Integrated Program in Humane Studies is Kenyon’s oldest interdisciplinary program. NEH Distinguished Teaching Professorship; Senior Faculty Trustee Teaching Excellence Award; selected by the senior class to deliver the 2024 Baccalaureate Address.

- Mentored 400+ student research projects whose publications have been downloaded more than 107,000 times, reaching 198 countries.

- Increased participation among underrepresented groups: 61% women, 13% Black, and 11% Latinx researchers.
- Five program alumni now hold AI-safety research roles at leading organizations.

Key Courses Created

Programming Humanity (IPHS 200) — introduction to programming for the humanities and social sciences within an ethically aware framework (Python, data analysis, NLP, network analysis); now among Kenyon’s largest courses.

AI for Humanity (IPHS 300) — upper-division applied machine learning, ethics, and social-science applications, covering GOFAI, deep neural networks, transformers, and large language models.

Cultural Analytics (IPHS 290) — computational approaches to cultural interpretation, from literary texts to social media.

Senior Research Seminar (IPHS 484) — capstone research projects.

Storylab — leveraging emotional arcs for narrative analysis using the SentimentArcs methodology.

AI Strategy — six-week executive course bridging technical AI and organizational implementation, created and filmed at the Bloomberg NYC office (2024).

CREATIVE PRACTICE & RESIDENCIES

Artist-in-Residence, WPI Global Lab, Worcester Polytechnic Institute (2025). Public talk and podcast on AI at the crossroads of governance, creativity, and the humanities.

DivaBot (2021). Transformer-based AI performing live improvisation with actress Lauren Katz (Thymele Arts, Los Angeles). Created with Jim Dennen (Wexner Center Artist-in-Residence) and Jon Chun for the centennial of Čapek’s *R.U.R.*, the play that introduced the word “robot.”

SELECTED PUBLIC ENGAGEMENT & MEDIA

Paul Baier, “Where AI Meets the Humanities: Inside Kenyon College’s Bold Experiment.” *Forbes* (November 2025).

Mitch Leslie, “AI’s Talent for Translation Lowers Language Barriers.” *Engineering* (Chinese Academy of Engineering / Elsevier) 55 (2025): 11–13.

“As AI Leaps Forward, Concern Rises That Innovation Is Leaving Safety Behind.” *Christian Science Monitor* (February 2026). Quoted as a NIST AI Safety Institute Consortium contributor.

“Could Artificial Intelligence Save Endangered Archives?” NPR / WOSU (February 2026). Lead subject, Archival Intelligence project.

Susan D’Agostino, “AI Raises Complicated Questions About Authorship.” *Inside Higher Ed* (2023).

MLA Member Spotlight (2025); MLA newsletter feature on the consortium’s NIST work (2026).

“Is AI Better at Making Art Than Humans?” AI Jazeera “The Stream” debate with Boris Eldagsen and Shane Balkowitsch (2023).

“Kenyon Professors Discuss Implications of ChatGPT, AI Art.” *Kenyon Collegian* (March 2023).

Bloomberg: AI Strategy Course, industry expert and on-screen presenter (2024).

Podcasts: “From Homer to GPT: The Collision of Human Imagination and AI,” *Humanity at Scale* with Bruce Temkin (2025); *Creative Velocity: AI Deep Dive* with Leslie Grandy (2026); “Translating Worlds: AI Expert’s Adventure from Pages to Pixels,” *Merging Minds* (2024); The Jarmarl Thomas Podcast, “AI and Creativity” (2025); *RadioAI*, “ChatGPT and Large Language Models” (2023).

EDITORIAL & REVIEWING SERVICE

Editorial Roles

- Editor and Contributor, *Oxford Studies in Philosophy and Literature* (Proust volume, 2022).
- Editorial Board Member, *International Journal of Literature and Arts*.
- Guest Editor, “Depiction of Good and Evil in Fairytales,” *Humanities* (MDPI), 2024–25.

Conference Program Committees & Reviewing

- Program Committee, AAAI/ACM Conference on AI, Ethics, and Society (AIES) (2025).
- Reviewer, ICML (2025, 2026, including the Position Paper Track) and NeurIPS (2026).

Journal & Grant Reviewing

Has completed 35+ peer-review reports (2021–2026). Journals reviewed include: PNAS; Science Advances; Nature Scientific Reports; npj Artificial Intelligence; AI & Society; Artificial Intelligence Review; Artificial Intelligence; Minds and Machines; Philosophia; American Philosophical Quarterly (AI Ethics special issue); Humanities and Social Sciences Communications; Politics and Governance; IEEE Transactions on Technology and Society; Behaviour & Information Technology; ACM Journal on Computing and Cultural Heritage; Journal of Cultural Analytics; Computational Humanities Research; Digital Humanities Quarterly; International Journal of Digital Humanities; International Journal of Humanities and Arts Computing; International Journal of Social Research Methodology; The Journal of Interactive Technology and Pedagogy; Poetics Today; Modernism/modernity; Texas Studies in Literature and Language; Interdisciplinary Studies in Literature and the Environment (ISLE); Essays in Economic & Business History; Frontiers in Psychology; Discover Artificial Intelligence; Discover Applied Sciences; Discover Education; Minerva; and Oxford University Press. Grant and research-program review for Schmidt Sciences (Trustworthy AI), the National Endowment for the Humanities, the Swiss National Science Foundation, and Maryland Industrial Partnerships (MIPS).

Tenure, Promotion & Program Review

- External promotion/tenure reviewer, Olin College of Engineering (2023).
- External program evaluator, English, Gustavus Adolphus College (2026).
- Computer Science Faculty Review, Harvey Mudd College.
- External Curriculum Review, Minerva College.

TECHNICAL EXPERTISE

Nearly a decade of hands-on implementation (2016–present). Fine-tuned and evaluated models from GPT-2 through contemporary LLMs including GPT-4, Claude, Llama, and Mistral. Python, TensorFlow, PyTorch, scikit-learn, HuggingFace Transformers. Red-teaming, threat modeling, vulnerability detection, explainable AI (XAI), uncertainty quantification, and confidence-aware processing for high-stakes applications. Developed ensemble methods for diachronic sentiment analysis (the methodology behind *The Shapes of Stories*; SentimentArcs is one implementation). Developed the Syntactic Framing Fragility framework and the Syntactic Variation Index (SVI), the first operational method for quantifying LLM decision consistency under logically equivalent syntactic transformations, aligned to EU AI Act and NIST risk categories. Multi-agent architecture design and evaluation; data visualization (Matplotlib, Seaborn); network analysis (NetworkX); SQL and Pandas; API integration for model evaluation; historical and multilingual computational analysis; topic modeling and stylometry.

LANGUAGES

French, German, Spanish (professional proficiency; French near-native). Latin, Ancient Greek (reading proficiency).